

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2005, 08:40:21 (without alignments)  
433.851 Million cell updates/sec

Title: US-10-695-140-2

Perfect score: 947 Gapext 0.5

Sequence: 1 APITAYSQQTGLIGCIITS.....RGVAKAVDFVPVESMETTMR 180

Scoring table: BLOSUM62

Post-processing: Maximum Match 100%

Total number of hits satisfying chosen parameters: 181204

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Maximum Match 100%

Listing First 45 summaries

Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubcaa/US07\_PUBCOMB.pep:\*

2: /cgn2\_6/ptodata/1/pubcaa/US06\_PUBCOMB.pep:\*

3: /cgn2\_6/ptodata/1/pubcaa/US10\_PUBCOMB.pep:\*

4: /cgn2\_6/ptodata/1/pubcaa/US07\_NEW\_PUB.pep:\*

5: /cgn2\_6/ptodata/1/pubcaa/PTCUS\_PUBCOMB.pep:\*

6: /cgn2\_6/ptodata/1/pubcaa/US08\_PUBCOMB.pep:\*

7: /cgn2\_6/ptodata/1/pubcaa/US08\_PUBCOMB.pep:\*

8: /cgn2\_6/ptodata/1/pubcaa/US09A\_PUBCOMB.pep:\*

9: /cgn2\_6/ptodata/1/pubcaa/US09B\_PUBCOMB.pep:\*

10: /cgn2\_6/ptodata/1/pubcaa/US09C\_PUBCOMB.pep:\*

11: /cgn2\_6/ptodata/1/pubcaa/US09 NEW\_PUB.pep:\*

12: /cgn2\_6/ptodata/1/pubcaa/US10\_PUBCOMB.pep:\*

13: /cgn2\_6/ptodata/1/pubcaa/US10A\_PUBCOMB.pep:\*

14: /cgn2\_6/ptodata/1/pubcaa/US10B\_PUBCOMB.pep:\*

15: /cgn2\_6/ptodata/1/pubcaa/US10C\_PUBCOMB.pep:\*

16: /cgn2\_6/ptodata/1/pubcaa/US10D\_PUBCOMB.pep:\*

17: /cgn2\_6/ptodata/1/pubcaa/US10E\_PUBCOMB.pep:\*

18: /cgn2\_6/ptodata/1/pubcaa/US10 NEW\_PUB.pep:\*

19: /cgn2\_6/ptodata/1/pubcaa/US11A\_PUBCOMB.pep:\*

20: /cgn2\_6/ptodata/1/pubcaa/US11 NEW\_PUB.pep:\*

21: /cgn2\_6/ptodata/1/pubcaa/US60\_NEW\_PUB.pep:\*

22: /cgn2\_6/ptodata/1/pubcaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	947	100.0	180	17	Sequence 2, Appli
2	947	100.0	685	18	Sequence 2, Appli
3	947	100.0	1985	14	Sequence 4,2, Appli
4	947	100.0	1985	16	Sequence 2, Appli
5	947	100.0	1985	20	Sequence 4,2, Appli
6	947	100.0	2201	13	Sequence 3, Appli
7	947	100.0	2201	14	Sequence 3, Appli
8	947	100.0	2201	16	Sequence 3, Appli
9	947	100.0	3010	15	Sequence 1, Appli
10	947	100.0	3010	15	Sequence 1, Appli
11	99.7		1692	14	Sequence 18, Appli

RESULT 1  
US-10-695-140-2  
; Sequence 2, Application US/10695140  
; Publication No. US2005001975A1  
; GENERAL INFORMATION:  
; APPLICANT: Boehringer Ingelheim International GmbH  
; TITLE OF INVENTION: Inhibitor-Resistant HCV NS3 Protease  
; FILE REFERENCE: 13/1090US  
; CURRENT APPLICATION NUMBER: US/10/695,140  
; CURRENT FILING DATE: 2003-10-28  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: HCV peptide  
; US-10-695-140-2

Query Match : 100.0%; Score 947; DB 17; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.e-92;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APITAYSQQTGLIGCITSLSLGRDNRQVGEVQVVSATOSPLATCYNQVCTTYHAGG 60  
Db 1 APITAYSQQTGLIGCITSLSLGRDNRQVGEVQVVSATOSPLATCYNQVCTTYHAGG 60  
Qy 61 SKTLAGPKGPITONMYNDLIVQWQPPGARSITPCTCGSSDPLYLVTRHADVTPVRRG 120  
Db 61 SKTLAGPKGPITONMYNDLIVQWQPPGARSITPCTCGSSDPLYLVTRHADVTPVRRG 120  
Qy 121 DSSQSLSPRPVSYLKGSGGPGLLCPSSHAVGIFRAAVCTRGVYKAVDVFVPEMETMR 180  
Db 121 DSRSQSLSPRPVSYLKGSGGPGLLCPSSHAVGIFRAAVCTRGVYKAVDVFVPEMETMR 180

#### ALIGNMENTS

GenCore version 5.1.6  
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2005, 08:37:01 ; Search time 22 Seconds

(without alignments)  
 610.765 Million cell updates/sec

Title: US-10-695-140-2

Perfect score: 947  
 Sequence: 1 APIAYSQQTRLGILGCITS.....RGYAKAVDFVFPVSEMETTMR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing First 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*

6: /cgn2\_6/ptodata/1/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	947	100.0	1985	4	US-09-539-601-9	Sequence 9, Appli
2	947	100.0	1985	4	US-09-539-601-12	Sequence 12, Appli
3	947	100.0	1985	4	US-09-539-601-18	Sequence 18, Appli
4	947	100.0	1985	4	US-09-539-601-4	Sequence 24, Appli
5	947	100.0	2201	4	US-09-539-601-6	Sequence 6, Appli
6	947	100.0	2201	4	US-09-539-601-15	Sequence 15, Appli
7	947	100.0	2201	4	US-10-029-907-3	Sequence 3, Appli
8	947	100.0	3010	4	US-09-539-601-3	Sequence 3, Appli
9	947	100.0	3010	4	US-09-539-601-21	Sequence 21, Appli
10	947	100.0	3010	4	US-09-539-601-77	Sequence 21, Appli
11	944	99.7	181	3	US-09-031-961-4	Sequence 4, Appli
12	944	99.7	631	2	US-08-936-665-1	Sequence 1, Appli
13	944	99.7	632	4	US-09-198-723A-23	Sequence 23, Appli
14	944	99.7	646	3	US-09-198-723A-69	Sequence 69, Appli
15	944	99.7	646	4	US-09-198-723A-72	Sequence 72, Appli
16	944	99.7	646	4	US-09-684-881-69	Sequence 69, Appli
17	944	99.7	646	4	US-09-684-881-72	Sequence 72, Appli
18	944	99.7	1692	3	US-09-263-933-4	Sequence 4, Appli
19	944	99.7	1692	4	US-09-919-901-4	Sequence 4, Appli
20	944	99.7	1692	4	US-10-191-966-4	Sequence 4, Appli
21	944	99.7	1692	4	US-10-191-966-4	Sequence 2, Appli
22	944	99.7	2307	3	US-09-263-933-2	Sequence 2, Appli
23	944	99.7	2307	4	US-09-919-901-2	Sequence 2, Appli
24	944	99.7	2307	4	US-10-191-966-2	Sequence 66, Appli
25	941	99.4	646	4	US-09-684-881-66	Sequence 66, Appli
26	941	99.4	646	4	US-09-684-881-66	Sequence 4, Appli
27	941	99.4	1692	3	US-09-263-933-11	Sequence 11, Appli

#### ALIGNMENTS

28	941	99.4	1692	3	US-09-263-933-18	Sequence 18, Appli
29	941	99.4	1692	4	US-09-919-901-11	Sequence 11, Appli
30	941	99.4	1692	4	US-09-919-901-18	Sequence 11, Appli
31	941	99.4	1692	4	US-10-191-966-11	Sequence 18, Appli
32	941	99.4	1692	4	US-10-191-966-18	Sequence 9, Appli
33	941	99.4	2307	3	US-09-263-933-9	Sequence 16, Appli
34	941	99.4	2307	4	US-09-919-901-9	Sequence 9, Appli
35	941	99.4	2307	4	US-09-919-901-16	Sequence 16, Appli
36	941	99.4	2307	4	US-10-191-966-9	Sequence 16, Appli
37	941	99.4	2307	4	US-10-191-966-16	Sequence 16, Appli
38	941	99.4	2307	4	US-10-191-966-16	Sequence 30, Appli
39	940	99.3	1985	4	US-09-539-601-30	Sequence 15, Appli
40	940	99.3	3010	4	US-10-017-736C-15	Sequence 10, Appli
41	937	98.9	292	4	US-10-017-736C-10	Sequence 16, Appli
42	937	98.9	303	4	US-10-017-736C-16	Sequence 16, Appli
43	937	98.9	303	4	US-10-017-736C-18	Sequence 18, Appli
44	937	98.9	334	4	US-10-017-736C-4	Sequence 4, Appli

RESULT 1						
US-09-539-601-9						
; Sequence 9, Application US/09539601C						
; Patent No. 6630343						
; GENERAL INFORMATION:						
; APPLICANT: Bartschenschlager, Ralf FW						
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System						
; FILE REFERENCE: all sequences						
; CURRENT APPLICATION NUMBER: US/09/539,601C						
; CURRENT FILING DATE: 2001-08-30						
; EARLIER APPLICATION NUMBER: 1999-04-03						
; EARLIER FILING DATE: 1999-04-03						
; NUMBER OF SEQ ID NOS: 51						
; SOFTWARE: PatentIn Ver. 2.1						
; SEQ ID NO: 9						
; LENGTH: 1985						
; TYPE: PRT						
; ORGANISM: Hepatitis C virus						
; US-09-539-601-9						
; Query						
; Match						
; Best Local Similarity						
; Matches 180; Conservative 100.0%; Mismatches 0; Indels 0; Gaps 0;						
; RESULT 2						
; US-09-539-601-12						
; Sequence 12, Appli						
; Sequence 18, Appli						
; Sequence 24, Appli						
; Sequence 6, Appli						
; Sequence 15, Appli						
; Sequence 3, Appli						
; Sequence 21, Appli						
; Sequence 2, Appli						
; Sequence 4, Appli						
; Sequence 23, Appli						
; Sequence 69, Appli						
; Sequence 72, Appli						
; Sequence 4, Appli						
; Sequence 2, Appli						
; Sequence 6, Appli						
; Sequence 1, Appli						
; Sequence 2, Appli						
; Sequence 3, Appli						
; Sequence 23, Appli						
; Sequence 4, Appli						
; Sequence 5, Appli						
; Sequence 6, Appli						
; Sequence 7, Appli						
; Sequence 8, Appli						
; Sequence 9, Appli						
; Sequence 10, Appli						
; Sequence 11, Appli						
; Sequence 12, Appli						
; Sequence 13, Appli						
; Sequence 14, Appli						
; Sequence 15, Appli						
; Sequence 16, Appli						
; Sequence 17, Appli						
; Sequence 18, Appli						
; Sequence 19, Appli						
; Sequence 20, Appli						
; Sequence 21, Appli						
; Sequence 22, Appli						
; Sequence 23, Appli						
; Sequence 24, Appli						
; Sequence 25, Appli						
; Sequence 26, Appli						
; Sequence 27, Appli						
; Sequence 28, Appli						
; Sequence 29, Appli						
; Sequence 30, Appli						
; Sequence 31, Appli						
; Sequence 32, Appli						
; Sequence 33, Appli						
; Sequence 34, Appli						
; Sequence 35, Appli						
; Sequence 36, Appli						
; Sequence 37, Appli						
; Sequence 38, Appli						

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model!

Run on: September 15, 2005, 08:34:45 ; Search time 177 Seconds  
(without alignments)  
520.758 Million cell updates/sec

Title: US-10-695-140-2

Perfect score: 947

Sequence: 1 APITAYSQOTRLGLGCITS.....RGVAKAVDFVPEVSMETTMR 180

Scoring table: BLOSUM62

Gappen 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : UniProt 03:  
1: uniprot\_sprot:  
2: uniprot\_trembl:  
\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	947	100.0	3010	2	Q9WMX2		Q9wmx2 hepatitis C
2	944	99.7	180	2	Q8QW30		Q8Qw30 hepatitis C
3	944	99.7	3010	2	P8A803		P8A803 hepatitis C
4	944	99.7	3010	2	P90191		P90191 hepatitis C
5	943	99.6	3010	2	Q9AU29		Q9au29 hepatitis C
6	943	99.6	180	2	Q8QVX9		Q8Qvx9 hepatitis C
7	943	99.6	181	2	Q91RP8		Q91rp8 hepatitis C
8	943	99.6	1984	2	Q7T4V8		Q7t4v8 hepatitis C
9	943	99.6	3010	2	Q9DTF0		Q9dtf0 hepatitis C
10	941	99.4	180	2	Q8BM13		Q8bm13 hepatitis C
11	941	99.4	3010	2	Q6Y9Y9		Q6y9y9 hepatitis C
12	940	99.4	3010	2	Q9DTD9		Q9dtD9 hepatitis C
13	940	99.3	180	2	Q8BLV3		Q8blv3 hepatitis C
14	940	99.3	180	2	Q80M29		Q80m29 hepatitis C
15	940	99.3	180	2	Q80M32		Q80m32 hepatitis C
16	940	99.3	181	2	Q91RQ4		Q91rq4 hepatitis C
17	940	99.3	234	2	Q70823		Q70823 hepatitis C
18	940	99.3	234	2	Q70824		Q70824 hepatitis C
19	940	99.3	234	2	Q70825		Q70825 hepatitis C
20	940	99.3	234	2	Q70827		Q70827 hepatitis C
21	940	99.3	234	2	Q70829		Q70829 hepatitis C
22	940	99.3	234	2	Q70830		Q70830 hepatitis C
23	940	99.3	234	2	Q70831		Q70831 hepatitis C
24	940	99.3	234	2	Q70832		Q70832 hepatitis C
25	940	99.3	234	2	Q89253		Q89253 hepatitis C
26	940	99.3	234	2	Q89254		Q89254 hepatitis C
27	940	99.3	234	2	Q89255		Q89255 hepatitis C
28	940	99.3	234	2	Q89256		Q89256 hepatitis C
29	940	99.3	361	2	Q70818		Q70818 hepatitis C
30	940	99.3	3008	2	Q9j3F4		Q9j3f4 hepatitis C
31	940	99.3	3010	2	Q68533		Q68533 hepatitis C

## ALIGNMENTS

RESULT 1									
ID	Q9WMX2	PRELIMINARY;	PRT;	3010 AA.					
AC	Q9WMX2;						SEQUENCE FROM N.A.		
DT	01-NOV-1999 (TREMBLrel. 12, Created)						STRAIN=RB;		
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)						MEDLINE=9932193; PubMed=10390360; DOI=10.1126/science.285.5424.110;		
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)						Lohmann V., Koerner F., Koch J.O., Herian U., Theilmann L., Bartschschager R.;		
DE	Polyprotein.						Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line."		
OS	Heptatitis C virus type 1b.						Science 285:110-113(1999); EMBL; Aut23879; CAB46677.i; -.		
RA	RP						PIR; A61196; A61196.		
RA	RP						PIR; P00246; P00246.		
RA	RP						HSSP; Q8BYXSL1; ICWX.		
RA	RP						DR; GO; 0016021; C:integral to membrane; IEA.		
RA	RP						DR; GO; 0019028; C:virion capsid; IEA.		
RA	RP						DR; GO; 0019031; C:virial envelope; IEA.		
RA	RP						DR; GO; 0005524; F:ATP binding; IEA.		
RA	RP						DR; GO; 0008726; F:RNA binding; IEA.		
RA	RP						DR; GO; 0003722; F:RNA binding; IEA.		
RA	RP						DR; GO; 0003968; F:RNA-directed RNA polymerase activity; IEA.		
RA	RP						DR; GO; 0008236; F:serine-type peptidase activity; IEA.		
RA	RP						DR; GO; 0005198; F:structural molecule activity; IEA.		
RA	RP						DR; GO; 0006508; P:proteolysis and peptidolysis; IEA.		
RA	RP						DR; GO; 0006350; P:transcription; IEA.		
RA	RP						DR; GO; 0019079; P:viral genome replication; IEA.		
RA	RP						DR; GO; 00019087; P:viral transformation; IEA.		
RA	RP						DR; InterPro; IPR000345; CycC_heme_BS.		
RA	RP						DR; InterPro; IPR014110; DEAD/DEAH_N.		
RA	RP						DR; InterPro; IPR011545; DEAD/DEAH_N.		
RA	RP						DR; InterPro; IPR002522; HCV_capsid.		
RA	RP						DR; InterPro; IPR002521; HCV_core.		
RA	RP						DR; InterPro; IPR002519; HCV_env.		
RA	RP						DR; InterPro; IPR002531; HCV_NS4.		
RA	RP						DR; InterPro; IPR000745; HCV_NS4a.		
RA	RP						DR; InterPro; IPR01490; HCV_NS4b.		
RA	RP						DR; InterPro; IPR02868; HCV_NS5a.		
RA	RP						DR; InterPro; IPR02166; HCV_RdRP.		
RA	RP						DR; InterPro; IPR001650; HeliTcase_C.		
RA	RP						DR; InterPro; IPR04109; Peptide_id_S20.		
RA	RP						DR; InterPro; IPR009003; Pept_Ser_Cys.		
RA	RP						DR; InterPro; IPR002518; Pept_U39_HCV_NS2.		

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: September 15, 2005, 08:35:26 ; Search time 39 Seconds

Title: US-10-695-140-2  
Perfect score: 947  
Sequence: 1 APITAYSQQTGQLLGLITS.....RGVAKAVDFVVEVESMETTMR 180

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : PIR 79.1

1: Pir1;\*  
2: Pir2;\*  
3: Pir3;\*  
4: Pir4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	926	97.8	3010	1 A45573	Genome polyprotein
2	923	97.5	3010	1 GNWVCJ	Genome polyprotein
3	921	97.3	3010	1 GNWVTC	Genome polyprotein
4	918	96.9	3010	1 GNWVTW	Genome polyprotein
5	904	95.5	3010	1 S18030	Genome polyprotein
6	884	93.3	3011	1 S13121	Genome polyprotein
7	882	93.1	3011	1 GNWVC3	Genome polyprotein
8	879	92.8	3011	1 GNWVCH	Genome polyprotein
9	797	84.2	3014	1 JCS620	Genome polyprotein
10	722	76.2	3033	1 GNWVJB	Genome polyprotein
11	720	76.0	3033	1 JQ1303	Genome polyprotein
12	276	29.1	3005	2 T08841	polyprotein - dour
13	256.5	27.1	2970	2 T08839	polyprotein - dour
14	86.5	9.1	398	2 B11284	probable periplasm
15	84	8.9	209	2 H83144	probable aromatic
16	84	8.9	904	2 A84212	hypothetical prote
17	81.5	8.6	786	2 S22155	oncogene 1 ('tre-2
18	78	8.4	233	2 T35594	hypothetical prote
19	79.5	8.4	868	2 H81775	aconitase hydratase
20	79	8.3	739	2 B88553	protein K04H4.2b [
21	79	8.3	3414	1 GRWVNB	genome polyprotein
22	78.5	8.3	1236	2 E70977	hypothetical prote
23	78	8.2	452	2 I39383	angio associated m
24	78	8.2	3412	1 GRWVTB	genome polyprotein
25	76.5	8.1	495	2 B11360	hypothetical prote
26	75.5	8.0	264	2 I38136	chymotrypsin-like
27	75	7.9	1089	2 S22158	htrA-like serine p
28	75	7.9	433	2 H97199	transforming prote
29	75	7.9	1399	2 G83112	DNA-directed RNA p

Query Match 97.8%; Score 926; DB 1; Length 3010;  
Best Local Similarity 96.1%; Pred. No. 4.6e-77; Mismatches 5; Indels 0; Gaps 0;

#### DESCRIPTION

Genome polyprotein

#### RESULT 1

A45573

Genome polyprotein - hepatitis C virus (strain JT)

N; Contains: capsid protein C; envelope protein M; hepatitis C virus

protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C; Species: hepatitis C virus

C; Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004

C; Accession: A45573

R; Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, Virus Res. 23, 39-53, 1992

A; Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier

A; Reference number: A45573; MUID:92299714; PMID:1318627

A; Accession: A45573

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-3010 <TAN>

A; Cross-references: UNIPROT:000269; GB:D11168; PIDN:9221612; PIDN:BA01943.1

A; Experimental source: HCV-UTP

A; Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBI:106207)

C; Superfamily: hepatitis C virus genome polyprotein

C; Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; ser:

F; 116-19-/Product: capsid protein C #status predicted <EM>

F; 116-389/Product: major envelope protein M #status predicted <EM>

F; 116-19-/Product: nonstructural protein NS1 #status predicted <NS1>

F; 130-728/Product: nonstructural protein NS2 #status predicted <NS2>

F; 130-1068/Product: nonstructural protein NS3 #status predicted <NS3>

F; 1007-1615/Product: hepatitis C virus genome polyprotein

F; 1230-1237/Region: nucleotide-binding motif A (P-loop)

F; 1312-1319/Region: DEXH motif B

F; 1516-1662/Product: nonstructural protein NS4a #status predicted <NA>

F; 1853-2013/Product: nonstructural protein NS4b #status predicted <NS5>

F; 2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match 97.8%; Score 926; DB 1; Length 3010;

Best Local Similarity 96.1%; Pred. No. 4.6e-77;

Mismatches 5; Indels 0; Gaps 0;

Matches 173; Conservative 5;

Score 926;

DB 1;

Length 3010;

QY 1 APITAYSOQTGQLIGCITSLTGRDRNOVEGEVQVQVSTATOSFLATCNGVYCWTYYHGAG 60

Db 1027 APITAYAQQTGQLLICIVSLTGRDKAQVQEVSQVQVSTATOSFLATCNGVYCWTYYHGAG 1086

QY 61 SKTLGPKGPKITOMNTNVDLGQAPPGARSLSITPCTGSSSDL1LYTREDAVDIVYRRG 120

Db 1087 SKTLGPKGPKITOMNTNVDLGQAPPGARSLSITPCTGSSSDL1LYTREDAVDIVYRRG 1146

QY 121 DSRGSLSPPPVPSKKGSSGPILLPSGSHAVGIFAAVTRGVAKAVDFVYVSEMETTMR 180

Db 1147 DGRGSLSPRSPVSYLKGSSEGPLLCPSGHAVGIFAAVTRGVAKAVDFVYVSEMETTMR 1206

#### RESULT 2

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: September 15, 2005, 08:34:26 ; Search time 72 Seconds  
(without alignments)  
966.901 Million cell updates/sec

Title: US-10-695-140-2  
Perfect score: 947  
Sequence: 1 APITAYSQQTQRGLGCCITS.....RGYAKAVDFVPVEMETM 180

Scoring table: BLOSUM62

Gapext 10.0 , Gapext 0.5

Searched: 2105622 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:  
1: geneseqP1980s: \*  
2: geneseqP1990s: \*  
3: geneseqP2000s: \*  
4: geneseqP2001s: \*  
5: geneseqP2002s: \*  
6: geneseqP2003s: \*  
7: geneseqP2003bs: \*  
8: geneseqP2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	947	100.0	180	8	ADN96855		Adn96855 Hepatitis
2	947	100.0	1985	5	AA018001		AA018001 Hepatitis
3	947	100.0	1985	5	AAE15729		AAE15729 Hepatitis
4	947	100.0	1985	5	AAE15731		AAE15731 Hepatitis
5	947	100.0	1985	5	AAE15717		AAE15717 Hepatitis
6	947	100.0	1985	5	AAE15727		AAE15727 Hepatitis
7	947	100.0	1985	5	AAE15728		AAE15728 Hepatitis
8	947	100.0	1985	5	AAE15722		AAE15722 Hepatitis
9	947	100.0	1985	5	AAE15730		AAE15730 Hepatitis
10	947	100.0	1985	8	ADJ57846		Adj57846 HCV Epi
11	947	100.0	2201	5	ABG30601		Abg30601 Hepatitis
12	947	100.0	2201	5	ABG30591		Abg30591 Hepatitis
13	947	100.0	2201	5	ABG30600		Abg30600 Hepatitis
14	947	100.0	2201	5	ABG30581		Abg30581 Hepatitis
15	947	100.0	2201	5	ABG30593		Abg30593 Hepatitis
16	947	100.0	2201	5	ABG30582		Abg30582 Hepatitis
17	947	100.0	2201	5	ABG30580		Abg30580 Hepatitis
18	947	100.0	2201	5	ABG30587		Abg30587 Hepatitis
19	947	100.0	2201	5	ABG30599		Abg30599 Hepatitis
20	947	100.0	2201	5	ABG30594		Abg30594 Hepatitis
21	947	100.0	2201	5	ABG30598		Abg30598 Hepatitis
22	947	100.0	2201	5	ABG30595		Abg30595 Hepatitis
23	947	100.0	3010	5	ABG32458		Abg32458 Hepatitis
24	947	100.0	3010	5	ABG32459		Abg32459 Hepatitis
25	947	100.0	3010	5	ABG32451		Abg32451 Hepatitis

#### ALIGNMENTS

RESULT 1  
ID ADN96855 standard; protein; 180 AA.

AC ADN96855;  
XX  
XX 29-JUL-2004 (first entry)  
DT XX Hepatitis C virus NS3 protease.  
DE XX NS3 protease; mutation; protease inhibitor; enzyme.  
KW XX Hepatitis C virus.  
XX 40 944 99.7 1985 8 ADN96855 Hepatitis  
41 944 99.7 2201 5 ABG30589 Hepatitis  
42 944 99.7 2201 5 ABG30583 Hepatitis  
43 944 99.7 2201 5 ABG30588 Hepatitis  
44 944 99.7 2307 3 AAY70054 Recombina  
45 943 99.6 1985 5 AAE15720 Hepatitis

XX  
XX WO2004039970-A1.  
XX 29-OCT-2002; 2002US-0421943P.  
XX (BOEHINGER INGELHEIM INT GMBH.  
XX Kukolj G, Lagace L, Marquis M, Lamarre D, Pause A;  
XX PI DR WPI; 2004-400164/37.  
XX N-PSDB; ADN96854.  
XX PT Novel hepatitis C virus non-structural 3 protease, useful for identifying potential second generation inhibitor of hepatitis C virus non-structural 3 protease.  
PT PR XX 24-OCT-2003; 2003WO-CR001636.  
XX PD 13-MAY-2004.  
XX PP 24-OCT-2003; 2003WO-CR001636.  
XX PR XX 29-OCT-2002; 2002US-0421943P.

XX (BOEHINGER INGELHEIM INT GMBH.

XX The invention relates to a hepatitis C virus (HCV) NS3 protease (I), or an amino acid sequence that is 90% identical to (I), comprising one or more mutations in at least one amino acid position Arg155Gln, Arg155Trp, Ala156Gly, Ala156Thr, Ala156Val. (I) is useful for identifying a potential second generation inhibitor of HCV NS3 protease activity which involves incubating an inhibitor-resistant (I) in the presence or absence of a candidate second generation inhibitor compound, and measuring the protease activity of the inhibitor-resistant (I) in the presence and absence of the candidate second generation inhibitor compound, where a decrease in the activity of the protease in the presence of a candidate second generation inhibitor indicates that the compound inhibits the hepatitis C virus non-structural 3 protease. (II) containing the gene encoding (I) are useful for evaluating HCV NS3 protease activity of inhibitor-